meaningful DATA

Making your data **worthwhile.**

LEI into SDMX with open-source tools

Antonio Olleros

Eurofiling Conference, 3rd June 2025

Aim of the presentation

Talking about standards

Discussing about metadata-driven systems

SDMX and VTL Open-Source solutions as enablers to metadata-driven systems

Standards

Short intro to LEI, SDMX and VTL

What is the LEI, and why should you be interested?

What is the LEI?

- A worldwide public **unique identifier** for companies (**who is who**)…
- And beyond! (**who owns whom**)?

Why is it relevant?

- Increased **transparency and trust** between legal entities globally
- Streamlined **counterparty due-diligence** and KYC processes, facilitating global trade
- Harmonized, centralized and **supplemented national business registry data**
- **Supports compliance** with regulatory requirements

What is SDMX, and why should you be interested?

What is SDMX?	**Standard** for **data and metadata** management and exchange
	Equivalent to DPM + Serialisation format (XBRL) + API specification
Why does it matter?	**Supported and adopted** by many **regulators** worldwide
	Designed to cover **the full statistical process**
	Provides **standardised API** specification
	Promotes **open-source** solutions (check sdmx.io)

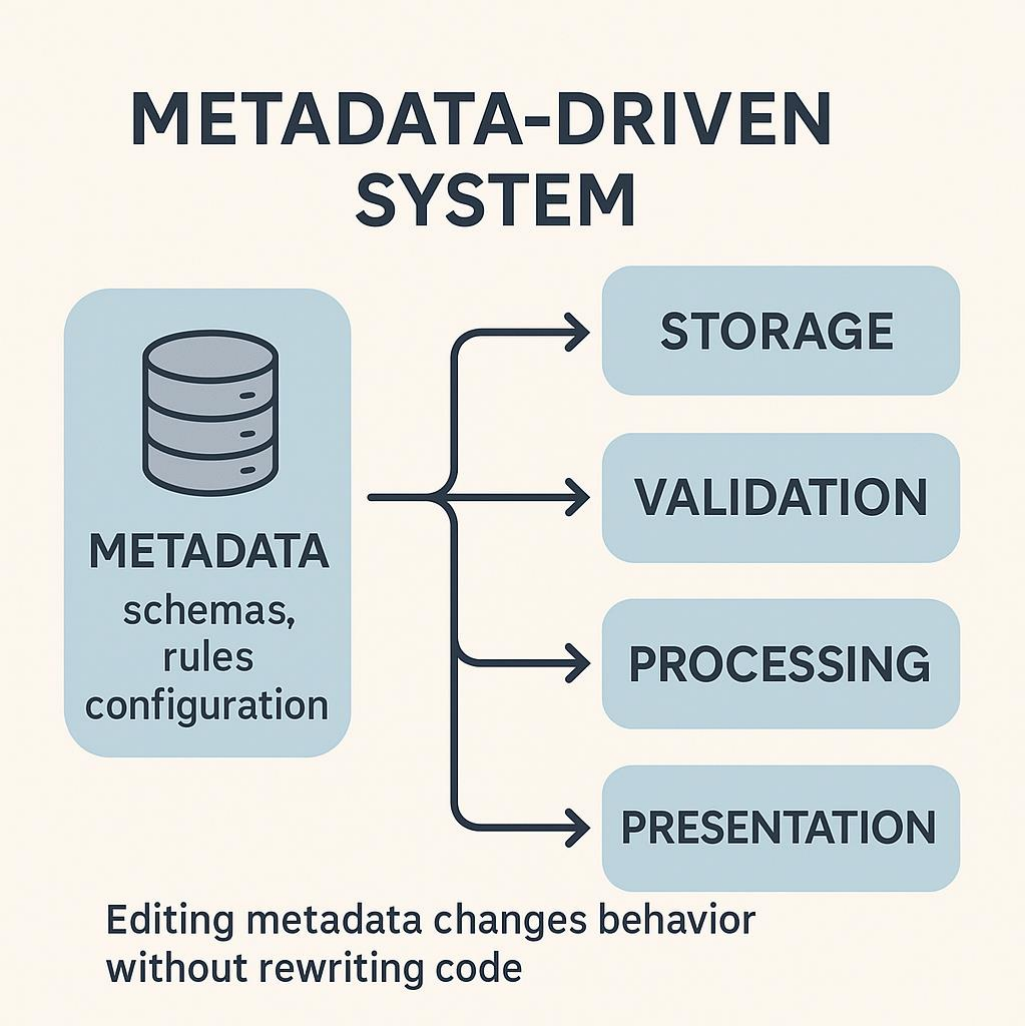
What is VTL, and why should you be interested?

- **Standard** for writing **validation and transformations**
- **Business user-oriented**
- Formal ➔ **Executable**
- **Technology agnostic**
- Very good compatibility with **SDMX**
- But compatible also with **other technologies** (e.g., DPM)

Metadata-driven systems

Focus on regulatory reporting/official statistics systems

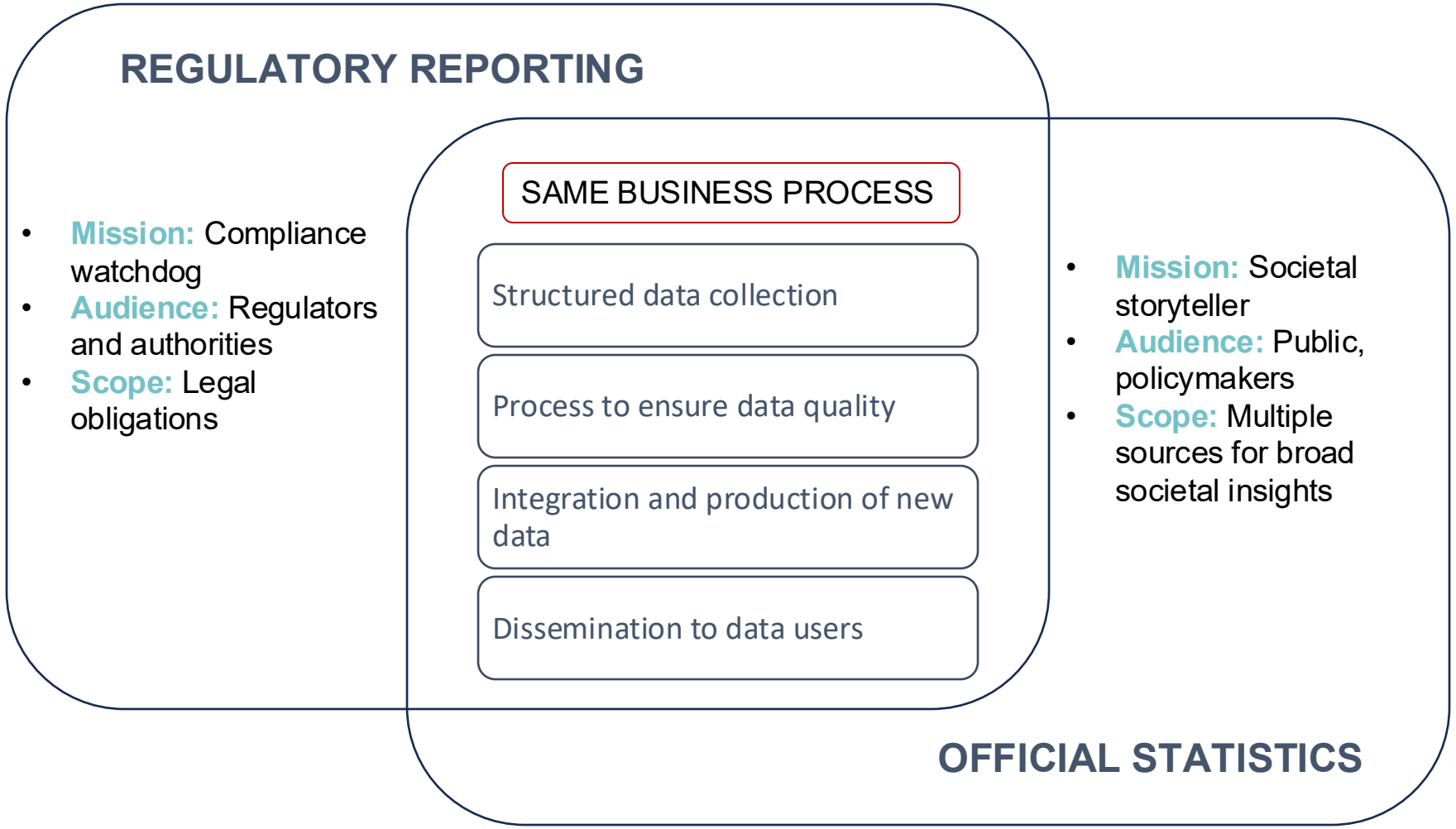
What is a metadata-driven system?



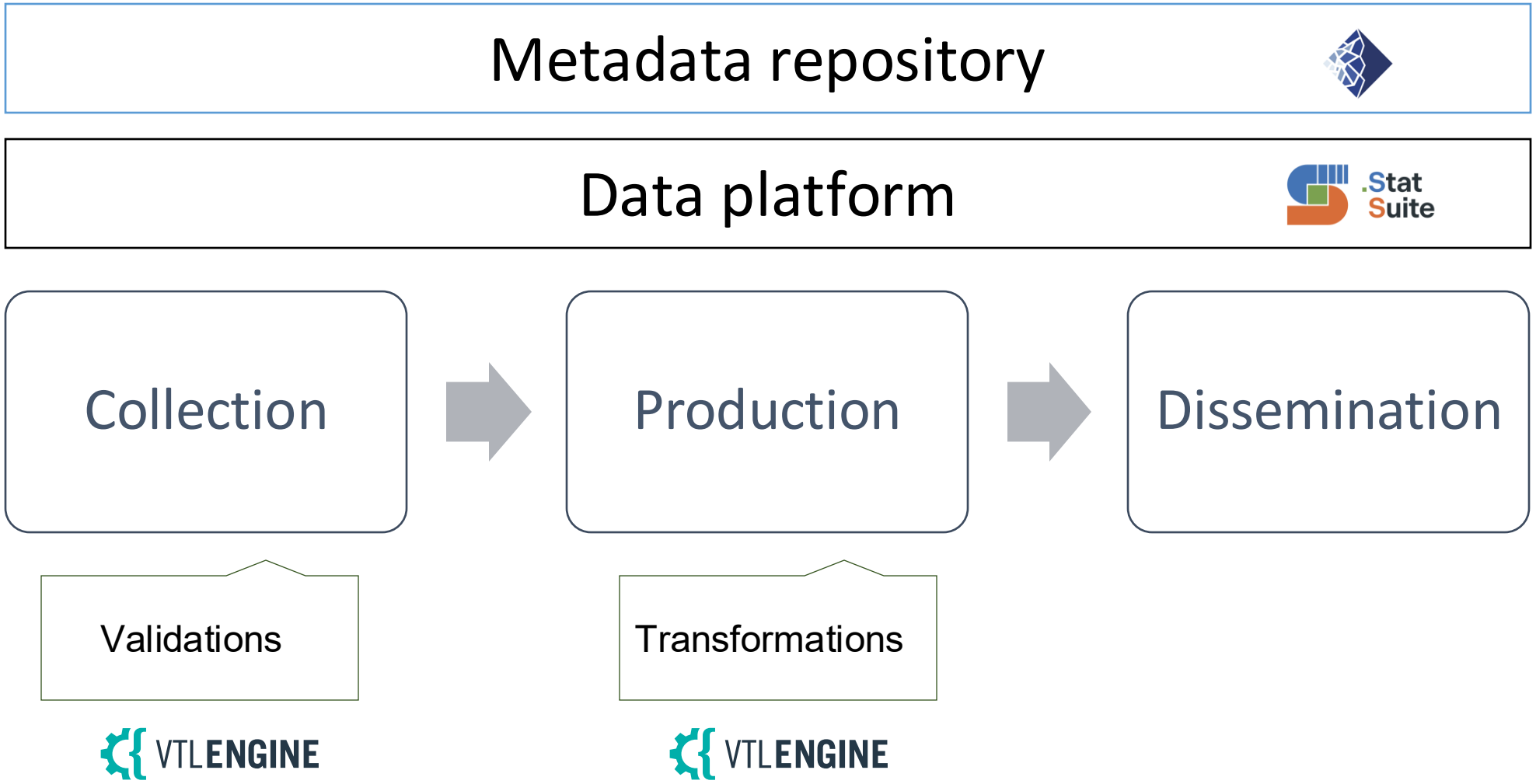
Disentangle IT development from business logic

Develop once, collect as many datasets as needed

Regulatory reporting or official statistics?



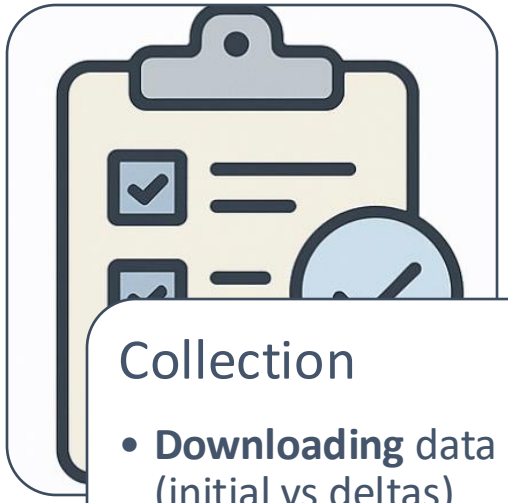
Metadata-driven statistical (regulatory reporting) system with open-source solutions



Example: LEI into SDMX realm

Focus on regulatory reporting/official statistics systems

The statistical process for the LEI case

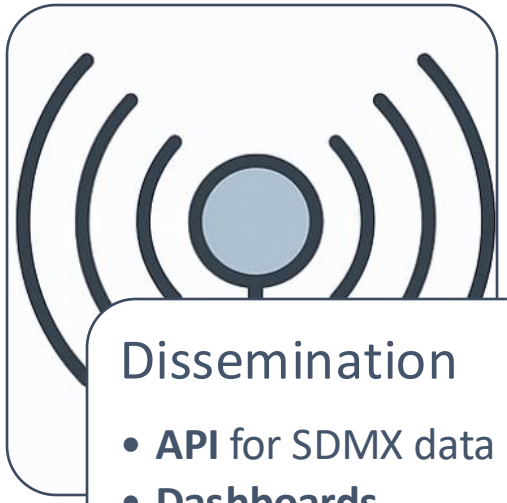


Collection

- **Downloading** data (initial vs deltas)
- **Converting** to SDMX
- **Structural validation**
- **Ingestion**

Production

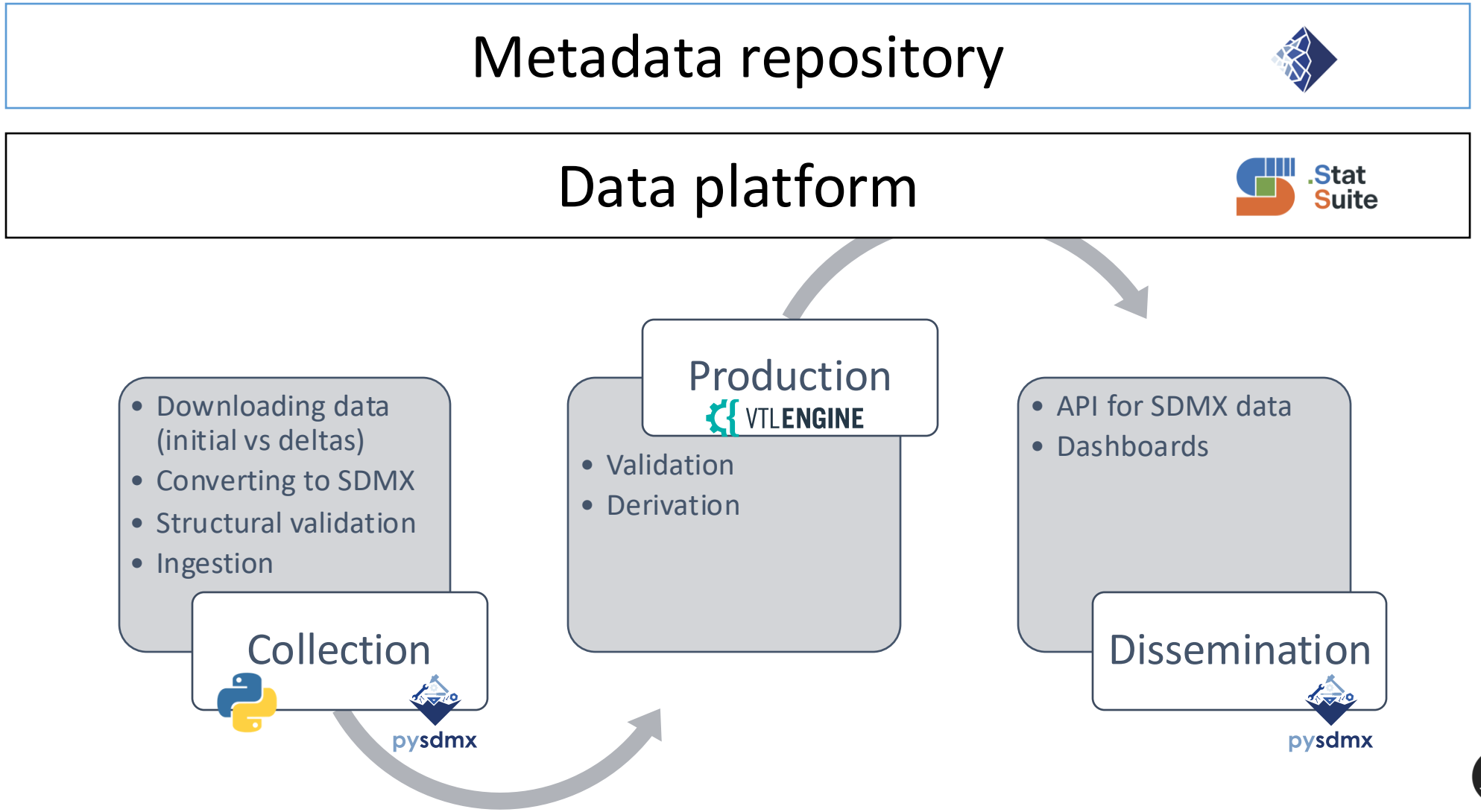
- **Validation**
- **Derivation**, eg:
  - Country aggregates
  - New variables
  - Analytical data



Dissemination

- **API** for SDMX data
- **Dashboards**

An open-source pipeline for LEI in SDMX



GitHub

SDMX snapshots

METADA DEFINITION

Viewing: LEI data [1.0]

Dimensions

[LEI] LEI

Measures

[POSTAL_CODE] Postal Code

[COUNTRY_INCORPORATION] COUNTRY_INCORPORATION

[COUNTRY_HEADQUARTERS] COUNTRY_HEADQUARTERS

[CATEGORY] CATEGORY

[SUBCATEGORY] SUBCATEGORY

[LEGAL_FORM] LEGAL_FORM

Dataset Attributes

- n/a -

Series Attributes

[LEGAL_NAME] LEGAL_NAME

Observation Attributes

- n/a -

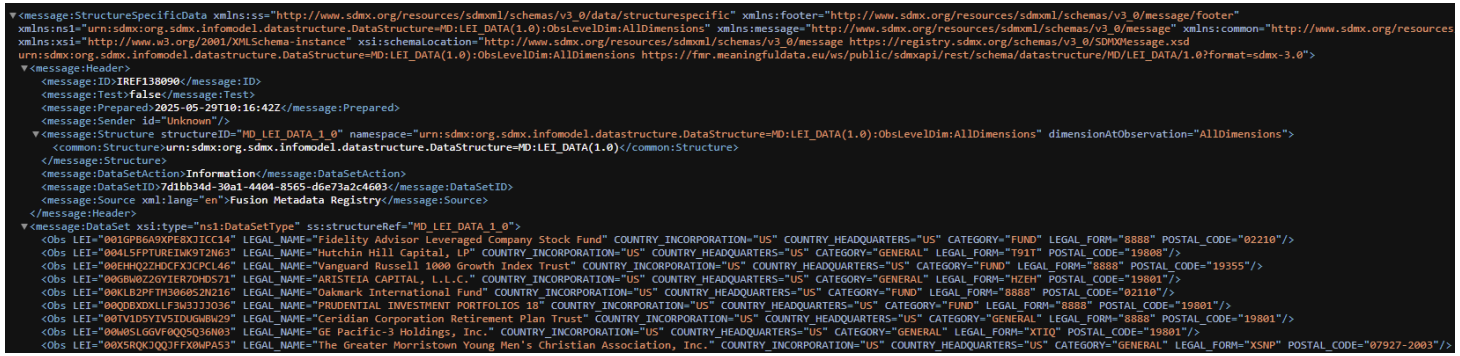
Group Attributes

- n/a -

SDMX-CSV file

STRUCTURI	STRUCTURI	ACTION	LEI	LEGAL_NAI	COUNTRY_	COUNTRY_	CATEGORY	SUBCATEG	LEGAL_FOI	POSTAL_CODE
datastructu	MD:LEI_DA	I	001GPB6A9XPE8XJICC14	Fidelity Adv	US	US	FUND		8888	02210
datastructu	MD:LEI_DA	I	004L5FPTUREIWK9T2N63	Hutchin Hil	US	US	GENERAL		T91T	19808
datastructu	MD:LEI_DA	I	00EHHQ2ZHDCFXJCPCL46	Vanguard R	US	US	FUND		8888	19355
datastructu	MD:LEI_DA	I	00GBW0Z2GYIER7DHDS71	ARISTEIA C/	US	US	GENERAL		HZEH	19801
datastructu	MD:LEI_DA	I	00KLB2PFTM3060S2N216	Oakmark In	US	US	FUND		8888	02110
datastructu	MD:LEI_DA	I	00ODBXDXLLF3W3IJJO36	PRUDENTIA	US	US	FUND		8888	19801

SDMX-XML file

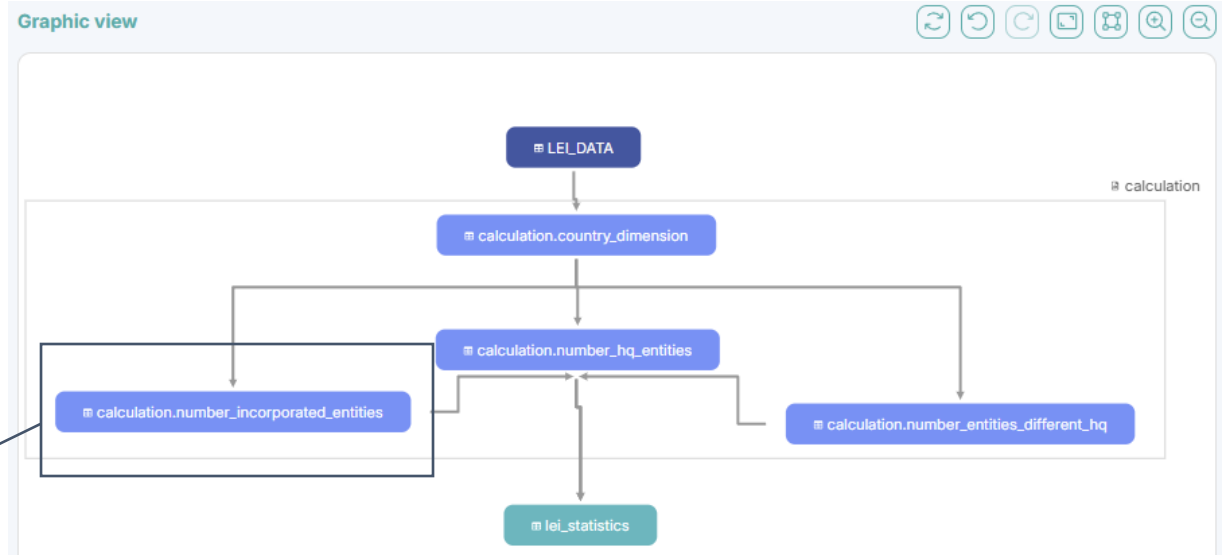


VTL snapshots

VALIDATIONS EXAMPLE

```
pc1: when COUNTRY_INCORPORATION = "AD" then match_characters(POSTAL_CODE,"^(AD\d{3})$") errorcode "pc1" errorlevel 3;
pc2: when COUNTRY_INCORPORATION = "AF" then match_characters(POSTAL_CODE,"^\d{4}$|^((([1-3][0-9])|(4[0-3]))\d{4})$") errorcod
pc3: when COUNTRY_INCORPORATION = "AI" then match_characters(POSTAL_CODE,"^AI-2640$") errorcode "pc3" errorlevel 3;
pc4: when COUNTRY_INCORPORATION = "AL" then match_characters(POSTAL_CODE,"^\d{4}$") errorcode "pc4" errorlevel 3;
pc5: when COUNTRY_INCORPORATION = "AM" then match_characters(POSTAL_CODE,"(\d{4})|(\d{6})") errorcode "pc5" errorlevel 3;
pc6: when COUNTRY_INCORPORATION = "AR" then match_characters(POSTAL_CODE,"^(([A-Z]\d{4}[A-Z]{3})|([A-Z]\d{4}))$") errorcode "
pc7: when COUNTRY_INCORPORATION = "AS" then match_characters(POSTAL_CODE,"^(96799(-\d{4})?)$") errorcode "pc7" errorlevel 3;
```

CALCULATIONS GRAPHICAL REPRESENTATION



ZOOMING ONE CALCULATION

```
number_incorporated_entities :=
    count(country_dimension#LEI group by COUNTRY_INCORPORATION)
        [calc identifier TYPE_MEASURE := "NUMBER_INCORPORATED_ENTITIES"]
        [rename int_var to OBS_VALUE, COUNTRY_INCORPORATION to COUNTRY];
```

meaningful DATA

Making your data **worthwhile.**

Thank you for your attention!

meaningfuldata.eu

Antonio Olleros

Founder & CEO

+34 645 89 16 57

antonio.olleros@meaningfuldata.eu